X is continuous $\rightarrow$ $\underbrace{P(X \le x)}_{\text{cdf}} = \int_{-\infty}^{x} f(x)\,dx$ density function: $\boxed{f(x) = \underbrace{c}_{\text{normalizing constant}} e^{-d(x)}}$

Continuous Distributions

A normal distribution and other density functions involving exponential forms play the most important role in probability and statistics. They are related in a certain way, as summarized in a diagram later in this topic.

Exponential distribution.

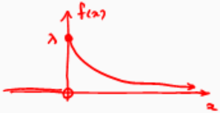
$\lambda > 0$

$$f(x) = \underset{\substack{\uparrow \\ \text{normalizing constant}}}{\lambda}\, e^{-\lambda x} \longleftarrow \int_0^\infty \lambda e^{-\lambda x}\,dx = \left[-e^{-\lambda x}\right]_0^\infty = \underbrace{\lim_{x\to\infty}(-e^{-\lambda x})}_{0} - \underbrace{(-e^{-\lambda(0)})}_{1}$$

$$= 1$$

The exponential density function is defined as

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0; \\ 0 & x < 0 \end{cases}$$

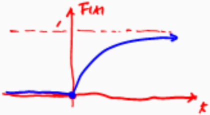


Then the cdf is computed as

$$F(t) = \int_0^t \lambda e^{-\lambda x}\,dx = \left[-e^{-\lambda x}\right]_0^t$$

$$= -e^{-\lambda t} - (-e^{-\lambda(0)})$$

$$= -e^{-\lambda t} + 1$$

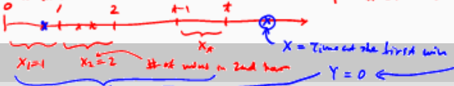


$$F(t) = \begin{cases} 1 - e^{-\lambda t} & t \geq 0; \\ 0 & t < 0 \end{cases}$$

where $\lambda > 0$ is called a **rate parameter**.

X_1 = # of wins in 1st hour $\quad E[X] = \lambda$

Y = # of wins in t hours $= \underbrace{X_1 + \cdots + X_t}_{\text{Poisson with}}$ $\quad E[X] = \lambda t$



Example

Suppose that the number of wins in a slot machine for an hour is distributed with Poisson distribution with λ. (a) If you play for t hours, what is the distribution for the number of wins in t hours? (b) Let X be the exact time of your first win. Find $P(X > t)$. (c) What is the distribution for X?

$$F(t) = P(X \le t) = 1 - P(X > t) = 1 - e^{-\lambda t} \xrightarrow{f(t) = \frac{d}{dt}F(t)} f(t) = \lambda e^{-\lambda t}$$

$$P(X > t) = P(Y = 0) = p(0) = e^{-\lambda t}\frac{(\lambda t)^0}{0!} = e^{-\lambda t}$$

# of wins in t hours

Poisson with λt

λ represent a rate of successes per unit of time

Example

Suppose that the number of wins in a slot machine for an hour is distributed with Poisson distribution with λ. (a) If you play for t hours, what is the distribution for the number of wins in t hours? (b) Let X be the exact time of your first win. Find $P(X > t)$. (c) What is the distribution for X?

(a) It is again a Poisson distribution with parameter λt. (b) Let Y be the number of wins in t hours. Then we have $P(X > t) = P(Y = 0)$. Thus,

$$P(X > t) = e^{-\lambda t} \frac{\lambda^0}{0!} = e^{-\lambda t}$$

(c) Since $F(t) = P(X \leq t) = 1 - P(X > t) = 1 - e^{-\lambda t}$, X must be exponentially distributed.

Survival function and memoryless property.

$F(t) = P(X \le t)$

$S(t) = 1 - \underbrace{F(t)}_{1-e^{-\lambda t}} = P(X > t)$

The function $S(t) = 1 - F(t) = P(X > t)$ is called the **survival function**. When $F(t)$ is an exponential distribution, we have $S(t) = e^{-\lambda t}$ for $t \ge 0$. Furthermore, we can find that

$$P(X > t + s \mid X > s) = \frac{P(X > t + s)}{P(X > s)} = \frac{S(t + s)}{S(s)} = S(t) = P(X > t), \tag{4.1}$$

lifetime

which is referred as the **memoryless** property of exponential distribution.

$$P(\underbrace{\{X > t+s\}}_{B} \mid \underbrace{\{X > s\}}_{A}) = P(B \mid A) = \frac{P(B \cap A)}{P(A)} = \frac{P(B)}{P(A)} = \frac{P(X > t+s)}{P(X > s)}$$

$$= \frac{S(t+s) = e^{-\lambda(t+s)}}{S(s) = e^{-\lambda s}} = e^{-\lambda t} = S(t) = P(X > t)$$

$B \cap A = B$

Example

A service time (in minutes) of post office is exponentially distributed with $\lambda = 0.2$. (a) Suppose that no one is there, and that you get the service immediately. Find the probability that it takes more than 5 minutes to complete the business. (b) Suppose that someone else is in service ahead of you, and that you are next in line. Find the probability that you wait more than 5 minutes.

(a)

0

X = time to complete the service ~ exponential with $\lambda = 0.2$

$$P(X > 5) = e^{-\lambda(5)} = e^{-1}$$

X = someone else's service time

(b)

5   5

Someone else started the service

You come in

Your service starts.

$$P(X > 5+5 \mid X > 5) = P(X > 5) = e^{-1}$$

Example

A service time (in minutes) of post office is exponentially distributed with $\lambda = 0.2$. (a) Suppose that no one is there, and that you get the service immediately. Find the probability that it takes more than 5 minutes to complete the business. (b) Suppose that someone else is in service ahead of you, and that you are next in line. Find the probability that you wait more than 5 minutes.

Let X be the service time. (a) We obtain

$$P(X > 5) = 1 - P(X \leq 5) = 1 - F(5) = e^{-1} \approx 0.37.$$

(b) The service of the person ahead of you started s minutes ago. $P(X > 5 + s | X > s) = P(X > 5) \approx 0.37.$

Goal: $Y = X_1 + \cdots + X_n$ As a special case $Y = X_1$

Exponential r.v.'s

The *gamma function*, denoted by $\Gamma(\alpha)$, is defined as

$$\Gamma(\alpha) = \int_0^\infty u^{\alpha-1} e^{-u}\, du, \quad \alpha > 0.$$

If $\alpha = 1$

$\Gamma(1) = \int_0^\infty e^{-u}\,du = 1$

$f(t) = \frac{\lambda^1}{\Gamma(1)} t^0 e^{-\lambda t} = \lambda e^{-\lambda t}$

exponential

Then the *gamma density* is defined as

$$f(t) = \frac{\lambda^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\lambda t} \quad t \geq 0;$$

$f(t) = 0$ if $t < 0$

which depends on two parameters $\alpha > 0$ and $\lambda > 0$. [The textbook uses a parameter $\beta = \frac{1}{\lambda}$ instead.]

Why do we need $\Gamma(\alpha)$?

Let $u = \lambda t$

$$\int_{-\infty}^{\infty} f(t)\,dt = \int_0^\infty \frac{\lambda^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\lambda t}\,dt = \frac{1}{\Gamma(\alpha)} \int_0^\infty \underbrace{(\lambda t)^{\alpha-1}}_{u^{\alpha-1}} \underbrace{e^{-\lambda t}}_{e^{-u}} \underbrace{\lambda\,dt}_{du} = \frac{1}{\Gamma(\alpha)} \underbrace{\int_0^\infty u^{\alpha-1} e^{-u}\,du}_{\Gamma(\alpha)} = 1$$

Properties of gamma function.

$$[-e^{-u}]_{u=0}^{u=\infty} = \underbrace{\lim_{u\to\infty}(-e^{-u})}_{0} - \underbrace{(-e^{-0})}_{1} = 1$$

$$\Gamma(1) \overset{\alpha=1}{=} \int_0^\infty e^{-u}\,du = 1$$

$$\Gamma(\alpha) = \int_0^\infty \underbrace{u^{\alpha-1}}_{v}\underbrace{e^{-u}\,du}_{dw} \quad \text{and } w = -e^{-u}$$

$$dv = (\alpha-1)u^{\alpha-2}$$

$$= \underbrace{\left[\underbrace{-e^{-u}}_{w}\underbrace{u^{\alpha-1}}_{v}\right]_{u=0}^{u=\infty}}_{0} + (\alpha-1)\underbrace{\int_0^\infty u^{\alpha-2}\underbrace{e^{-u}}_{-w}\,du}_{\Gamma(\beta)} \quad \beta = \alpha-1$$

$$= (\alpha-1)\Gamma(\alpha-1)$$

Suppose that the parameter α is a positive integer n. Then we have

$$\Gamma(n) \overset{\alpha=n}{=} (n-1)\Gamma(n-1) = (n-1)(n-2)\Gamma(n-2) = \cdots$$

$$= (n-1)\times\cdots\times 2\times 1\times \underbrace{\Gamma(1)}_{=1} = (n-1)!$$

The mgf of gamma distribution.

$E[e^{tx}] = \int_{-\infty}^{\infty} e^{tx} f(x)\, dx$

Let X be a gamma random variable with parameter (α, λ). For $t < \lambda$, the mgf $M_X(t)$ of X can be computed as follows.

$$M_X(t) = \int_0^\infty e^{tx} \overbrace{\frac{\lambda^\alpha}{\Gamma(\alpha)}}^{\text{constant}} x^{\alpha-1} e^{-\lambda x}\, dx = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{\alpha-1} e^{-(\lambda-t)x}\, dx$$

$$= \frac{\lambda^\alpha}{(\lambda-t)^\alpha} \frac{1}{\Gamma(\alpha)} \int_0^\infty (\underbrace{(\lambda-t)x}_{u})^{\alpha-1} e^{-\underbrace{(\lambda-t)x}_{u}} \underbrace{(\lambda-t)dx}_{du}$$

Let $u = (\lambda - t)x$

$du = (\lambda - t)dx$

$$= \left(\frac{\lambda}{\lambda-t}\right)^\alpha \frac{1}{\Gamma(\alpha)} \underbrace{\int_0^\infty u^{\alpha-1} e^{-u}\, du}_{\Gamma(\alpha)} = \left(\frac{\lambda}{\lambda-t}\right)^\alpha$$

If you use $\lambda = 1/\beta$ then we obtain $\boxed{M_X(t) = (1-\beta t)^{-\alpha}}$.

$\left(\frac{1/\beta}{1/\beta - t}\right)^\alpha = \left(\frac{1}{1-\beta t}\right)^\alpha = (1-\beta t)^{-\alpha}$

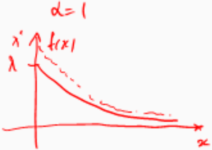
rate parameter per unit of time

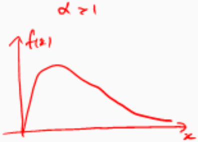
$\beta = \frac{1}{\lambda}$ = average time of event

$\lambda = 2$

average time of event $= \frac{1}{2}$

Parameters of gamma/exponential distribution.





We call the parameter α a **shape parameter**, because changing α changes the shape of the density. We call the parameter λ a **rate parameter**, because changing λ merely rescales the density without changing its shape. In particular, the gamma distribution with $\alpha = 1$ becomes an exponential distribution (with parameter λ).

$\lambda = \frac{1}{\beta}$

$\beta = \frac{1}{\lambda}$

The parameter $\beta = 1/\lambda$, if you choose, is called a **scale parameter**.

Property of scale parameter.

Let X be a service time (in minutes) having a gamma distribution. If we change the unit of measurement from minutes to seconds then $(60)X$ has a gamma distribution with new scale parameter $(60)\beta$. While the scale parameter changes from β to $(60)\beta$, the rate parameter changes from λ to $\lambda/60$.

Theorem

If X has a gamma distribution with parameters α and $\beta = 1/\lambda$ then $Y = kX$ has a gamma distribution with parameters α and $k\beta$.

Property of scale parameter.

Let X be a service time (in minutes) having a gamma distribution. If we change the unit of measurement from minutes to seconds then (60)X has a gamma distribution with new scale parameter (60)β. While the scale parameter changes from β to (60)β, the rate parameter changes from λ to $\lambda/60$.

Theorem

If X has a gamma distribution with parameters α and $\beta = 1/\lambda$ then $Y = kX$ has a gamma distribution with parameters α and $k\beta$. ← scale

rate $\frac{1}{k\beta} = \frac{\lambda}{k}$

We obtain

$= E[e^{tY}] = E[e^{(tk)X}]$

$$M_Y(t) = M_X(\underbrace{kt}_{t}) = [1 - \beta(kt)]^{-\alpha} = [1 - (\underbrace{k\beta}_{\beta'})t]^{-\alpha} = (1 - \beta' t)^{-\alpha}$$

Thus, Y has the mgf of gamma distribution with parameter α and $k\beta$.

rate parameter is $\frac{1}{k\beta} = \frac{\lambda}{k}$

Moments of gamma distribution.

$M_X(t) = (1-\beta t)^{-\alpha} \to \frac{d}{dt} M_X(t) = (-\beta)[-\alpha(1-\beta t)^{-\alpha-1}]$

$= \alpha\beta(1-\beta t)^{-(\alpha+1)}$

$M_X'(0) = \alpha\beta = \frac{\alpha}{\lambda}$

$\frac{d}{dt} M_X'(t) = \alpha\beta(-\beta)[-(\alpha+1)(1-\beta t)^{-\alpha-2}]$

$M_X''(0) = \alpha(\alpha+1)\beta^2 = \alpha(\alpha+1)\frac{1}{\lambda^2}$

We can find the first and second moment

$$E[X] = M_X'(0) = \frac{\alpha}{\lambda}$$

and

$$E[X^2] = M_X''(0) = \frac{\alpha(\alpha+1)}{\lambda^2}$$

Thus, we can obtain

$E[x^2] - (E[x])^2$

$$\mathrm{Var}(X) = \frac{\alpha(\alpha+1)}{\lambda^2} - \frac{\alpha^2}{\lambda^2} = \frac{\alpha}{\lambda^2}.$$

If you use $\beta = 1/\lambda$ then $E[X] = \alpha\beta$ and $\mathrm{Var}(X) = \alpha\beta^2$.

scale rate per unit of time

$\lambda = 2$

0

average time $= \frac{1}{2} = \beta$

Expectation by direct calculation.

$\int_{-\infty}^{\infty} x\, f(x)\, dx$

Let X be a gamma random variable with parameter (α, λ). Then we can calculate $E[X]$ and $E[X^2]$ as follows.

Let $u = \lambda x$

$$E[X] = \int_0^\infty x \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}\, dx = \frac{1}{\lambda \Gamma(\alpha)} \int_0^\infty (\underbrace{\lambda x}_{u})^\alpha e^{-\underbrace{\lambda x}_{u}}\, \underbrace{\lambda dx}_{du}$$

$$= \frac{1}{\lambda \Gamma(\alpha)} \underbrace{\int_0^\infty u^\alpha e^{-u}\, du}_{\Gamma(\alpha+1)} = \frac{1}{\lambda} \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)} = \frac{\alpha}{\lambda}$$

$\Gamma(\alpha+1) = \alpha\,\Gamma(\alpha)$

Variance by direct calculation.

We have

$\int_{-\infty}^{\infty} x^2 f(x)\,dx$

$$E[X^2] = \int_0^\infty x^2 \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}\,dx = \frac{1}{\lambda^2\Gamma(\alpha)} \int_0^\infty (\lambda x)^{\alpha+1} e^{-\lambda x}\,\lambda dx$$

$$= \frac{1}{\lambda^2\Gamma(\alpha)} \underbrace{\int_0^\infty u^{\alpha+1} e^{-u}\,du}_{\Gamma(\alpha+2)} = \frac{1}{\lambda^2}\frac{\Gamma(\alpha+2)}{\Gamma(\alpha)} = \frac{\alpha(\alpha+1)}{\lambda^2}$$

$\Gamma(\alpha+2) = (\alpha+1)\Gamma(\alpha+1)$
$= (\alpha+1)\alpha\Gamma(\alpha)$

Thus, we can obtain

$$\mathrm{Var}(X) = \frac{\alpha(\alpha+1)}{\lambda^2} - \frac{\alpha^2}{\lambda^2} = \frac{\alpha}{\lambda^2}$$

Sum of independent random variables.

Let X and Y be independent gamma random variables with the respective parameters (α_1, λ) and (α_2, λ). Then the sum $Z = X + Y$ of random variables has the mgf

$E[e^{(X+Y)t}] = E[e^{tX} \cdot e^{tY}]$

$$M_Z(t) = M_X(t) \cdot M_Y(t) = \left(\frac{\lambda}{\lambda - t}\right)^{\alpha_1} \times \left(\frac{\lambda}{\lambda - t}\right)^{\alpha_2} = \left(\frac{\lambda}{\lambda - t}\right)^{\alpha_1 + \alpha_2}$$

$M_X(t) = \left(\frac{\lambda}{\lambda - t}\right)^{\alpha_1}$, $M_Y(t) = \left(\frac{\lambda}{\lambda - t}\right)^{\alpha_2}$, $\alpha_1 + \alpha_2 = \alpha$

which is the mgf of gamma distribution with parameter $(\alpha_1 + \alpha_2, \lambda)$. Thus, Z is a gamma random variable with parameter $(\alpha_1 + \alpha_2, \lambda)$.

If $X_1, X_2, \dots, X_n$ are independent exponential r.v's — Gamma with $(\alpha = 1, \lambda)$

$$Y = X_1 + X_2 + \cdots + X_n = \underbrace{X_1 + X_2}_{\text{Gamma with } (\alpha=2,\lambda)} + \cdots + X_n \sim \text{Gamma with } (\alpha = n, \lambda)$$

X_1: Gamma with $(\alpha = 1, \lambda)$ = Exponential

Gamma with $(\alpha = 2, \lambda)$ ≠ Exponential

Chi-square distribution.

Another special case of gamma r.v. with $\alpha = \frac{n}{2}$, $\lambda = \frac{1}{2}$

It consists of n independent gamma r.v's $X_1, \dots, X_n$ of $\alpha = \frac{1}{2}$, $\lambda = \frac{1}{2}$

$Y = X_1 + \dots + X_n \sim$ gamma with $\alpha = \frac{n}{2}$, $\lambda = \frac{1}{2}$

A special case of gamma distribution associated with normal distribution instead of exponential

The gamma distribution

$\alpha = \frac{1}{2}, 1, \frac{3}{2}, 2, \dots$

$$f(x) = \frac{1}{2^{n/2}\Gamma(n/2)} x^{n/2-1} e^{-x/2}, \quad x \geq 0$$

with $\alpha = \frac{n}{2}$ and $\lambda = \frac{1}{2}$ is called the **chi-square distribution** with n degrees of freedom. [It plays a vital role later in understanding another important distribution, called t-distribution later.] A chi-square random variable X has the mean $E[X] = n$ and the variance $\mathrm{Var}(X) = 2n$.

$E[X] = \frac{\alpha}{\lambda} = \frac{n/2}{1/2} = n$

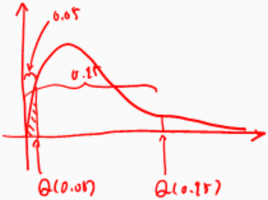
$\mathrm{Var}(X) = \frac{\alpha}{\lambda^2} = \frac{n/2}{(1/2)^2} = 2n$

Example

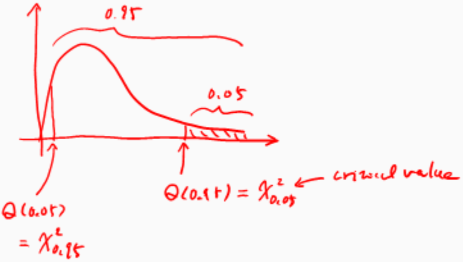
Let X be a chi-square random variable with $n = 6$.

1. Find the quantile $Q(0.05)$ and $Q(0.95)$.
2. Let Y be a gamma random variable with $\alpha = 3$ and $\lambda = 1/4$. Then find $P(3.28 < Y < 25.2)$.

Quantile



Critical value



$\frac{Y}{2} \sim$ Gamma with $\alpha = 3 = \frac{n}{2}$ and rate $\frac{\lambda}{1/2} = \frac{1/4}{1/2} = \frac{1}{2} \sim$ Chi-square with 6 degree of freedom

$k = \frac{1}{2}$

Example

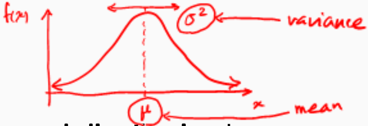
Let X be a chi-square random variable with $n = 6$.

1. Find the quantile $Q(0.05)$ and $Q(0.95)$.
2. Let Y be a gamma random variable with $\alpha = 3$ and $\lambda = 1/4$. Then find $P(3.28 < Y < 25.2)$.

1. From the table we obtain $Q(0.05) \approx \overset{\chi^2_{0.95}}{1.635}$ and $Q(0.95) \approx \overset{\chi^2_{0.05}}{12.592}$.
2. Since $X = Y/2$ has a chi-square random variable with $n = 6$, we obtain

$$P(\underset{2}{\underline{3.28}} < \underset{2}{\underline{Y}} < \underset{2}{\underline{25.2}}) = P(1.64 < X < 12.6) = \underbrace{F(12.6)}_{0.95} - \underbrace{F(1.64)}_{0.05}$$
$$\approx 0.95 - 0.05 = 0.9$$

Normal distribution.



A **normal distribution** is represented by a family of distributions which have the same general shape, sometimes described as "bell shaped." The normal distribution has the pdf

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad (4.2)$$

which depends upon two parameters μ and σ^2. In (4.2), $\pi = 3.14159\ldots$ is the famous "pi" (the ratio of the circumference of a circle to its diameter), and $\exp(u)$ is the exponential function e^u with the base $e = 2.71828\ldots$ of the natural logarithm.

Much celebrated integrations are the following:

optional problem

$$
\begin{aligned}
(1)\quad \int_{-\infty}^{\infty} f(x)\,dx &= \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx = 1. \qquad (4.3)\\
E[X] = \int_{-\infty}^{\infty} x f(x)\,dx &= \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} x e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx = \mu.\\
Var(X) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx &= \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} (x-\mu)^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx = \sigma^2.
\end{aligned}
$$

$$E[(X - \underbrace{E[X]}_{\mu})^2]$$

Equation (4.3) guarantees that the function (4.2) always represents a "probability density" no matter what values the parameters μ and σ^2 would take.

2. $$\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} x\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}(t+\mu)\, e^{-\frac{t^2}{2\sigma^2}}\,dt = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} t\, e^{-\frac{t^2}{2\sigma^2}}\,dt + \frac{\mu}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{t^2}{2\sigma^2}}\,dt$$

$$\begin{pmatrix} t = x-\mu \\ x = t+\mu \\ dx = dt \end{pmatrix}$$

$$= \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} \underbrace{t\, e^{-\frac{t^2}{2\sigma^2}}}_{\text{odd function}}\,dt + \mu \overbrace{\int_{-\infty}^{\infty} g(t)\,dt}^{1} = \mu$$

$$\underbrace{\text{odd function}}_{=\,0}$$

$$g(t) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{t^2}{2\sigma^2}}$$

normal density function with mean 0 and variance σ^2

3. $$\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}(x-\mu)^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} \underbrace{t}_{u}\, \underbrace{t\, e^{-\frac{t^2}{2\sigma^2}}\,dt}_{dv} \longrightarrow v = -\sigma^2 e^{-\frac{t^2}{2\sigma^2}}$$

$$= \underbrace{\frac{1}{\sigma\sqrt{2\pi}}\left[-\sigma^2 t\, e^{-\frac{t^2}{2\sigma^2}}\right]_{-\infty}^{\infty}}_{0} + \underbrace{\frac{\sigma^2}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{t^2}{2\sigma^2}}\,dt}_{=} = \sigma^2$$

$$(\sigma^2)\underbrace{\int_{-\infty}^{\infty} g(t)\,dt}_{1} = \sigma^2$$

$$g(t) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{t^2}{2\sigma^2}}$$ is again a normal density

Parameters of normal distribution.

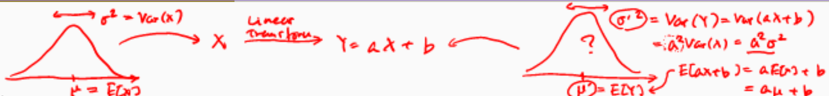
$\sqrt{Var(X)} = \sqrt{\sigma^2} = \sigma$

We say that a random variable X is **normally distributed** with parameter (μ, σ^2) when X has the pdf (4.2). Since $E[X] = \mu$ and $\mathrm{Var}(X) = \sigma^2$, the parameter μ is called a **mean** (or, location parameter), and the parameter σ a **standard deviation** (or, a scale parameter). The parameter μ provides the center of distribution, and the density function $f(x)$ is symmetric around μ. Small values of σ lead to high peaks but sharp drops, while larger values of σ lead to flatter densities. The shorthand notation

'distributed'

$$\boxed{X \sim N(\mu, \sigma^2)}$$

is often used to express that X is a normal random variable with parameter (μ, σ^2) with **variance** σ^2.



One of the important properties of normal distribution is that their linear transformation remains normal.

Theorem

Assume $a \neq 0$. If X is a normal random variable with parameter (μ, σ^2) [that is, the pdf of X is given by (4.2)], then $Y = aX + b$ is also a normal random variable having parameter $(\underbrace{a\mu + b}_{\text{mean}}, \underbrace{(a\sigma)^2}_{\text{variance}})$.

In particular,

$$\frac{X - \mu}{\sigma} = \underbrace{\left(\frac{1}{\sigma}\right)}_{a} X - \underbrace{\left(\frac{\mu}{\sigma}\right)}_{b} \sim N(\underbrace{a\mu+b}_{0}, \underbrace{(a\sigma)^2}_{1^2}) \quad (4.4)$$

becomes a normal random variable with parameter $(0, 1)$, called the **standard normal distribution**.

$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

$aX + b$

$=$

Assume that $a > 0$. Since X has the pdf (4.2), we obtain the cdf of Y as follows:

$P(Y \le t)$

$=$

$$F_Y(t) = P\left(aX + b \le t\right) = P\left(X \le \frac{t-b}{a}\right) = F_X\left(\frac{t-b}{a}\right)$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{x=-\infty}^{x=\frac{t-b}{a}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx$$

Let $\boxed{y = ax + b}$

$x = \frac{y-b}{a}$ $\quad dx = \frac{1}{a}\, dy$

$y = a\left(\frac{t-b}{a}\right) + b = t$

$-\frac{(x-\mu)^2}{2\sigma^2} = -\frac{\left[\frac{y-b}{a} - \mu\right]^2}{2\sigma^2} = \frac{\left[\frac{y-\mu a-b}{a}\right]^2}{2\sigma^2}$

$$= \frac{1}{a\sigma\sqrt{2\pi}} \int_{y=-\infty}^{y=t} e^{-\frac{(y-a\mu-b)^2}{2a^2\sigma^2}}\, dy = \int_{-\infty}^{t} f_Y(y)\, dy$$

Thus, Y has the pdf of $N(a\mu + b, a^2\sigma^2)$. The case with $a < 0$ is an exercise; see Problem 9.

$f_Y(y) = \frac{1}{a\sigma\sqrt{2\pi}} e^{-\frac{[y-(a\mu+b)]^2}{2(a\sigma)^2}} = \frac{1}{\sigma'\sqrt{2\pi}} e^{-\frac{[y-\mu']^2}{2(\sigma')^2}} \leftarrow$ It is a normal density

with $\mu' = a\mu + b$ and $\sigma' = a\sigma$

Standard normal distribution.

The normal density $\phi(x)$ with parameter $(\overset{\mu}{\overset{\shortparallel}{0}}, \overset{\sigma^2}{\overset{\shortparallel}{1}})$ is given by

$$\phi(x) := \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}},$$ standard normal density

and the **table of standard normal distribution** is used to obtain the values for the cdf

$$\Phi(t) := \int_{-\infty}^{t} \phi(x)\, dx.$$

How to calculate probabilities.

$X \sim N(\mu, \sigma^2)$

Suppose that a tomato plant height X is normally distributed with parameter (μ, σ^2). Then what is the probability that the tomato plant height is between a and b? The integration

$$P(a \leq X \leq b) = \int_a^b \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx$$

seems too complicated.

$X \sim N(\mu, \sigma^2)$

If we consider the random variable $\dfrac{X-\mu}{\sigma}$, then the event $P(\{a \le X \le b\})$ is equivalent to the event

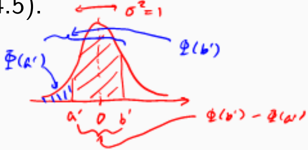
$Z \sim N(0,1)$

$$\left\{ \underbrace{\frac{a-\mu}{\sigma}}_{a'} \le \left(\frac{X-\mu}{\sigma}\right) \le \underbrace{\frac{b-\mu}{\sigma}}_{b'} \right\}.$$

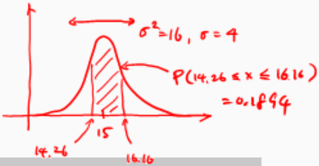
Let $a' = \frac{a-\mu}{\sigma}$ and $b' = \frac{b-\mu}{\sigma}$. Then in terms of probability, this means that

$$P(a \le X \le b) = P\left(a' \le \underbrace{\left(\frac{X-\mu}{\sigma}\right)}_{= Z} \le b'\right) = \int_{a'}^{b'} \phi(x)\,dx = \underbrace{\Phi(b') - \Phi(a')}. \tag{4.5}$$

Finally look at the values for $\Phi(a')$ and $\Phi(b')$ from the table of standard normal distribution, and plug them into (4.5).



$X \sim N(15, 16)$



Example

The tomato plant height is normally distributed with parameter $(15, 16)$ in inches. What is the probability that the height is between 14.24 and 16.16 inches?

$\mu = 15$ and $\sigma = 4$

$P(a \le X \le b) = P\left(\frac{a-\mu}{\sigma} \le \frac{x-\mu}{\sigma} \le \frac{b-\mu}{\sigma}\right)$

14.24 16.16

Example

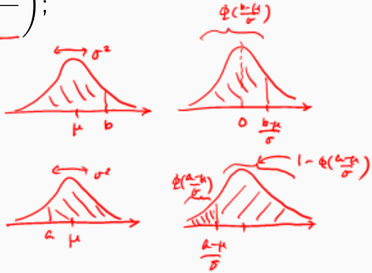
The tomato plant height is normally distributed with parameter $(15, 16)$ in inches. What is the probability that the height is between 14.24 and 16.16 inches?

We can calculate $a' = \frac{14.24-15}{4} = -0.19$ and $b' = \frac{16.16-15}{4} = 0.29$. Then find $\Phi(-0.19) = 0.4247$ and $\Phi(0.29) = 0.6141$. Thus, the probability of interest becomes 0.1894, or approximately 0.19.

Summary of probability calculations.

Suppose that a random variable X is normally distributed with mean μ and standard deviation σ. Then we can obtain

1. $P(a \le X \le b) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$;
2. $P(X \le b) = \Phi\left(\frac{b-\mu}{\sigma}\right)$;
3. $P(a \le X) = 1 - \Phi\left(\frac{a-\mu}{\sigma}\right)$.



The mgf of normal distribution.

$\mu=0$ and $\sigma^2=1$

$\sim N(0,1)$

$f(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$

Let X be a standard normal random variable. Then we can compute the mgf of X as follows.

$e^{-\frac{x^2}{2}+tx} = e^{-\frac{1}{2}\underbrace{(x^2-2xt+t^2)}_{(x-t)^2}+\frac{t^2}{2}} = e^{-\frac{(x-t)^2}{2}}\cdot e^{t^2/2}$

$$M_X(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{tx}\, e^{-\frac{x^2}{2}}\,dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{1}{2}(x-t)^2+\frac{t^2}{2}}\,dx$$

$= E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f(x)\,dx$

$$= e^{\frac{t^2}{2}}\underbrace{\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{(x-t)^2}{2}}\,dx}_{1} = \exp\left(\frac{t^2}{2}\right)$$

$\mu=t$

$\frac{1}{\sqrt{2\pi}}e^{-\frac{(x-t)^2}{2}} \sim N(t,1)$

Since $X=(Y-\mu)/\sigma$ becomes a standard normal random variable and $Y=\sigma X+\mu$, the mgf $M(t)$ of Y can be given by

$Y \sim N(\mu,\sigma^2)$

$$M(t) = e^{\mu t}M_X(\sigma t) = \underbrace{\exp\left(\frac{\sigma^2 t^2}{2}+\mu t\right)}_{\text{mgf for } Y}.$$

$= E[e^{tY}] = E[e^{t(\sigma X+\mu)}] = E[e^{t\sigma X}e^{t\mu}] = e^{t\mu}\,E[e^{(t\sigma)X}]$

Moments of normal distribution.

$X \sim N(\mu, \sigma^2)$

$M_X(t) = \exp\left[\frac{\sigma^2 t^2}{2} + \mu t\right] = e^{\frac{\sigma^2 t^2}{2} + \mu t}$

$M_X'(t) = [\sigma^2 t + \mu] e^{\frac{\sigma^2 t^2}{2} + \mu t}$

$M_X''(t) = \sigma^2 e^{\frac{\sigma^2 t^2}{2} + \mu t} + [\sigma^2 t + \mu]^2 e^{\frac{\sigma^2 t^2}{2} + \mu t}$

Then we can find the first and second moment

$$E[X] = M_X'(0) = \mu \quad \text{and} \quad E[X^2] = M''(0) = \underbrace{\mu^2 + \sigma^2}$$

Thus, we can obtain $\mathrm{Var}(X) = (\mu^2 + \sigma^2) - \mu^2 = \sigma^2$.

$= E[X^2] - (E[X])^2$

Sum of independent random variables

Let X and Y be independent normal random variables with the respective parameters (μ_x, σ_x^2) and (μ_y, σ_y^2). Then the sum $Z = X + Y$ of random variables has the mgf

$E[e^{tZ}] = E[e^{tX}]\, E[e^{tY}]$

$$M_Z(t) = M_X(t) \cdot M_Y(t) = \exp\left(\frac{\sigma_x^2 t^2}{2} + \mu_x t\right) \cdot \exp\left(\frac{\sigma_y^2 t^2}{2} + \mu_y t\right)$$

$$= \exp\left(\frac{(\sigma_x^2 + \sigma_y^2)t^2}{2} + (\mu_x + \mu_y)t\right) = \exp\left[\frac{\sigma^2 t^2}{2} + \mu t\right]$$

is the mgf of a normal r.v.

which is the mgf of normal distribution with parameter $(\mu, \sigma^2) = (\mu_x + \mu_y, \sigma_x^2 + \sigma_y^2)$. By the property (a) of mgf, we can find that Z is a normal random variable with parameter $(\mu_x + \mu_y, \sigma_x^2 + \sigma_y^2)$.

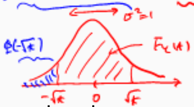
Z has a normal distribution. Then

$\mu = E[Z] = E[X] + E[Y] = \mu_x + \mu_y$

$\sigma^2 = Var(Z) = Var(X) + Var(Y) = \sigma_x^2 + \sigma_y^2$

$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$

Let X be a standard normal random variable, and let



$$Y = X^2 \ ? \longrightarrow \underset{P(Y \le t)}{F_Y(t)} = \int_0^t f_Y(y)\,dy \longrightarrow \frac{d}{dt}F_Y(t) = f_Y(t)$$

be the square of X. Then we can express the cdf F_Y of Y as $F_Y(t) = P(X^2 \le t) = P(-\sqrt{t} \le X \le \sqrt{t}) = \Phi(\sqrt{t}) - \Phi(-\sqrt{t})$ in terms of the standard normal cdf Φ. By differentiating the cdf, we can obtain the pdf f_Y as follows.

Gamma density $f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}$ ($\lambda = \frac{1}{2}$, $\alpha = \frac{1}{2} = \frac{n}{2}$ with $n=1$)

$$f_Y(t) = \frac{d}{dt}F_Y(t) = \underbrace{\frac{1}{\sqrt{2\pi}} t^{-\frac{1}{2}} e^{-\frac{t}{2}}}_{①} = \frac{1}{2^{1/2}\underbrace{\Gamma(1/2)}_{\sqrt{\pi}}} t^{-1/2} e^{-t/2},$$

where we use $\Gamma(1/2) = \sqrt{\pi}$. Thus, the square X^2 of X is the chi-square random variable with 1 degree of freedom.

$\Phi'(x) = \phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$

$$\frac{d}{dt}\left[\Phi(\sqrt{t}) - \Phi(-\sqrt{t})\right] = \frac{1}{2} t^{-1/2} \Phi'(\sqrt{t}) + \frac{1}{2} t^{-1/2} \Phi'(-\sqrt{t}) = \frac{1}{2} t^{-1/2} \frac{1}{\sqrt{2\pi}} e^{-t/2} + \frac{1}{2} t^{-1/2} \frac{1}{\sqrt{2\pi}} e^{-t/2} = ①$$

The relation, continued.

Now let $X_1, \ldots, X_n$ be independent and identically distributed random variables (iid random variables) of standard normal distribution. Since X_i^2's have the gamma distribution with parameter $(1/2, 1/2)$, the sum

χ^2 with $n=1$ $\longleftrightarrow$ Gamma with $\lambda = \frac{1}{2}$, $\alpha = \frac{1}{2}$

$$Y = \sum_{i=1}^{n} X_i^2 = \underbrace{X_1^2 + X_2^2}_{\text{Gamma with } \lambda=\frac{1}{2},\ \alpha=\frac{2}{2}=\frac{m}{2} \text{ with } m=2} + \cdots + X_n$$

Gamma $\longleftrightarrow$ χ^2 with $m=2$

has the gamma distribution with parameter $(\alpha = n/2, 1/2 = \lambda)$. That is, the sum Y has the chi-square distribution with n degree of freedom.

Extrema of exponential random variables.

possibly different rate parameters

Let $X_1, \dots, X_n$ be independent exponential random variables with the respective parameters $\lambda_1 \dots, \lambda_n$. Then we can define the random variable $V = \min(X_1, \dots, X_n)$ of the minimum of $X_1, \dots, X_n$. To find the distribution of V, consider the survival function $\underbrace{P(V > t)}_{1 - F_V(t)}$ of V and calculate as follows.

$A_k = \{X_k > t\}$

$$
\begin{aligned}
P(V > t) &= P(\{X_1 > t\} \cap \dots \cap \{X_n > t\}) = P(A_1 \cap A_2 \cap \dots \cap A_n) \\
&= P(X_1 > t) \times \dots \times P(X_n > t) \\
&= e^{-\lambda_1 t} \times \dots \times e^{-\lambda_1 t} = e^{-(\lambda_1 + \dots + \lambda_n)t},
\end{aligned}
$$

$F_{X_k}(t) = 1 - e^{-\lambda_k t}$

$P(X_k \geq t) = e^{-\lambda_k t}$

which is the survival function of exponential distribution with parameter $\lambda = \lambda_1 + \dots + \lambda_n$. Thus, V has an exponential distribution with parameter $\lambda = \lambda_1 + \dots + \lambda_n$.

$F_V(t) = 1 - e^{-(\lambda_1 + \dots + \lambda_n)t}$

$f_V(t) = F_V'(t) = (\lambda_1 + \dots + \lambda_n) e^{-(\lambda_1 + \dots + \lambda_n)t}$

$= \lambda e^{-\lambda t}$

Summary diagram.

Normal(μ, σ^2)

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty$$

$$E[X] = \mu, \quad Var(X) = \sigma^2$$

(1)

(2)

(3)

(5)

$E[$

Chi-square (with m degrees of freedom)

$f(x) =$

Relationship among random variables.

(1) If X_1 and X_2 are independent normal random variables with respective parameters (μ_1, σ_1^2) and (μ_2, σ_2^2), then $X = X_1 + X_2$ is a normal random variable with $(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

(2) If $X_1, \cdots, X_m$ are iid normal random variables with parameter (μ, σ^2), then $X = \sum_{k=1}^{m} \left(\frac{X_k - \mu}{\sigma} \right)^2$ is a chi-square (χ^2) random variable with m degrees of freedom.

(3) $\alpha = \frac{m}{2}$ and $\lambda = \frac{1}{2}$.

(4) If $X_1, \ldots, X_n$ are independent exponential random variables with respective parameters $\lambda_1, \cdots, \lambda_n$, then $X = \min\{X_1, \ldots, X_n\}$ is an exponential random variable with parameter $\sum_{i=1}^{n} \lambda_i$.

Relationship among random variables, continued.

(5) If $X_1, \cdots, X_n$ are iid exponential random variables with parameter λ, then $X = \sum_{k=1}^{n} X_k$ is a gamma random variable with $(\alpha = n, \lambda)$.

(6) $\alpha = 1$.

(7) If X_1 and X_2 are independent random variables having gamma distributions with respective parameters (α_1, λ) and (α_2, λ), then $X = X_1 + X_2$ is a gamma random variable with parameter $(\alpha_1 + \alpha_2, \lambda)$.

Exercises

Problem

Suppose that the lifetime of an electronic component follows an exponential distribution with rate parameter $\lambda = 0.2$*.*

1. *Find the probability that the lifetime is less than* 10*.*
2. *Find the probability that the lifetime is between* 5 *and* 15*.*

Problem

Suppose that in a certain population, individual's heights are approximately normally distributed with parameters $\mu = 70$ and $\sigma = 3in$.

1. *What proportion of the population is over 6ft. tall?*
2. *What is the distribution of heights if they are expressed in centimeters?*

X = Height in inches

Y = Height in cm = $(2.54)X$

Problem

Let X be a normal random variable with $\mu = 5$ and $\sigma = 10$.

1. *Find $P(X > 10)$.*
2. *Find $P(-20 < X < 15)$.*
3. *Find the value of x such that $P(X > x) = 0.05$.*

1. $P(|X-\mu| \le 0.675\sigma) = P(-0.675\sigma \le X-\mu \le 0.675\sigma)$

$= P\left(-0.675 \le \underbrace{\boxed{\frac{X-\mu}{\sigma}}}_{Z \sim N(0,1)} \le 0.675\right) = P(-0.675 \le Z \le 0.675) = \Phi(0.68) - \Phi(-0.67)$

Problem

Suppose that $X \sim N(\mu, \sigma^2)$.

1. *Find* $P(|X-\mu| \le 0.675\sigma)$.
2. *Find the value of* c *in terms of* σ *such that* $P(\mu - c \le X \le \mu + c) = 0.95$.

2. $P(\mu - c \le X \le \mu + c) = P\left(\frac{\mu-c-\mu}{\sigma} \le \underbrace{\frac{X-\mu}{\sigma}}_{Z \sim N(0,1)} \le \frac{\mu+c-\mu}{\sigma}\right)$

$= P\left(-\frac{c}{\sigma} \le Z \le \frac{c}{\sigma}\right) = 0.95$

$c = 1.96\sigma$

$0.975 = \Phi\left(\frac{c}{\sigma}\right)$

$0.975 = \Phi(z)$ for $z = 1.96$

Problem

X is an exponential random variable, and $P(X < 1) = 0.05$. What is λ?

Problem

The lifetime in months of each component in a microchip is exponentially distributed with parameter $\mu = \frac{1}{10000}$.

1. *For a particular component, we want to find the probability that the component breaks down before 5 months, and the probability that the component breaks down before 50 months. Use the approximation* $e^{-x} \approx 1 - x$ *for a small number* x *to calculate these probabilities.*
2. *Suppose that the microchip contains 1000 of such components and is operational until three of these components break down. The lifetimes of these components are independent and exponentially distributed with parameter* $\mu = \frac{1}{10000}$. *By using the Poisson approximation, find the probability that the microchip operates functionally for 5 months, and the probability that the microchip operates functionally for 50 months.*

X = Lifetime ~ Exponential with $\mu = \frac{1}{10{,}000}$

(a) $p = P(X \leq 5) = F_X(5) = 1 - e^{-\mu(5)} = 1 - e^{-0.0005} \approx 1 - (1 - 0.0005) = 0.0005$

$P(X \leq 50) = ?$

(b) Y = # of components breaking down out of $n = 1000$

~ Binomial with $(p = 0.0005, n = 1000)$

$\overset{\text{approx}}{\sim}$ Poisson with $\lambda = np = (1000)\frac{5}{10{,}000} = 0.5$

$P(Y \leq 2) = p_Y(0) + p_Y(1) + p_Y(2)$

Problem

The length of a phone call can be represented by an exponential distribution. Then answer to the following questions.

1. *A report says that 50% of phone call ends in 20 minutes. Assuming that this report is true, calculate the parameter λ for the exponential distribution.*
2. *Suppose that someone arrived immediately ahead of you at a public telephone booth. Using the result of (a), find the probability that you have to wait between* 10 *and* 20 *minutes.*

Problem

Let X_i*,* $i = 1, \ldots, 6$*, be independent standard normal random variables. Find*

$$P(X_1^2 + X_2^2 + X_3^2 + X_4^2 + X_5^2 + X_6^2 \leq 1.64)$$

Problem

Let X be a normal random variable with parameter (μ, σ^2), and let $Y = aX + b$ with $a < 0$.

1. *Show that*

$$P(Y \le t) = \int_{-\infty}^{t} f(x)\, dx,$$

where

$$f(x) = \frac{1}{|a|\sigma\sqrt{2\pi}} \exp\left[-\frac{(x - a\mu - b)^2}{2(a\sigma)^2}\right].$$

2. *Conclude that Y is a normal random variable with parameter $(a\mu + b, (a\sigma)^2)$.*

$$P(Y \le t) = P(aX + b \le t) = P(aX \le t - b) = P\left(X \ge \frac{t-b}{a}\right) = \int_{\frac{t-b}{a}}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

Let $u = ax + b$; $du = a\,dx = -|a|\,dx$

$$= \int_{-\infty}^{t} \frac{1}{|a|\sigma\sqrt{2\pi}} e^{-\frac{(u - a\mu - b)}{2a^2\sigma^2}} du$$

Problem

Let X be a normal random variable with parameter (μ, σ^2), and let $Y = e^X$. Find the density function f of Y by forming

$$P(Y \le t) = \int_0^t f(x)\, dx.$$

This density function f is called the lognormal density since $\ln Y$ *is normally distributed.*

$x = \ln y;\ y = e^x$

$$F_Y(t) = P(Y \le t) = P(e^x \le t) = P(x \le \ln t) = \int_{x=-\infty}^{x=\ln t} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx = \int_{y=0}^{y=t} f_Y(y)\, dy$$

Let $x = \ln y$; $dx = \frac{1}{y} dy$

$$= \int_{y=0}^{y=t} \underline{\frac{1}{y\sigma\sqrt{2\pi}} e^{-\frac{(\ln y-\mu)^2}{2\sigma^2}}}\, dy = \int_{y=0}^{y=t} \underline{f_Y(y)}\, dy$$

Optional problems

Problem

Show that the standard normal density integrates to unity by completing the following questions.

1. *Show that* $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{x^2+y^2}{2}}\,dx\,dy = \int_{0}^{\infty}\int_{0}^{2\pi} e^{-\frac{r^2}{2}}\,r\,d\theta\,dr = 2\pi.$
2. *Show that* $\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}\,dy = 1.$

Problem

Suppose that a non-negative random variable X has the memoryless property

$$P(X > t + s \mid X > s) = P(X > t).$$

By completing the following questions, we will show that X must be an exponential random variable.

1. *Let $S(t)$ is the survival function of X. Show that the memoryless property implies that $S(s + t) = S(s)S(t)$ for $s, t \geq 0$.*
2. *Let $\kappa = S(1)$. Show that $\kappa > 0$.*
3. *Show that $S\left(\frac{1}{n}\right) = \kappa^{\frac{1}{n}}$ and $S\left(\frac{m}{n}\right) = \kappa^{\frac{m}{n}}$.*

Therefore, we can obtain $S(t) = \kappa^t$ for any $t \geq 0$. By letting $\lambda = -\ln \kappa$, we can write $S(t) = e^{-\lambda t}$.

Problem

The gamma function is a generalized factorial function.

1. *Show* $\Gamma(\frac{1}{2}) = 2\int_0^\infty e^{-x^2}\,dx$.
   Hint: Use the substitution $u = x^2$.
2. *Show* $\int_{-\infty}^\infty e^{-x^2}\,dx = \sqrt{\pi}$.
   Hint: Use the result of optional problem 11.
3. *Show that* $\Gamma(\frac{1}{2}) = \sqrt{\pi}$, *and that*

$$\Gamma\left(\frac{n}{2}\right) = \frac{\sqrt{\pi}(n-1)!}{2^{n-1}\left(\frac{n-1}{2}\right)!}$$

   if n is an odd integer.

Answers to exercises

Problem 1.

Let X be the life time of the electronic component.

1. $P(X \leq 10) = 1 - e^{-(0.2)(10)} \approx 0.8647$
2. $P(5 \leq X \leq 15) = e^{-(0.2)(5)} - e^{-(0.2)(15)} \approx 0.3181$

Problem 2.

Let X be a normal random variable with $\mu = 70$ and $\sigma = 3$.

1. $P(X \geq (6)(12)) = 1 - \Phi\left(\frac{72-70}{3}\right) \approx 1 - \Phi(0.67) \approx 0.25$
2. In centimeters, the height becomes $(2.54)X$, and it is normally distributed with $\mu = (2.54)(70) = 177.8$ and $\sigma = (2.54)(3) = 7.62$.

Problem 3.

1. $P(X > 10) = 1 - \Phi\left(\frac{10-5}{10}\right) = 1 - \Phi(0.5) = 0.3085.$
2. $P(-20 < X < 15) = \Phi\left(\frac{15-5}{10}\right) - \Phi\left(\frac{-20-5}{10}\right) =$
   $\Phi(1) - \Phi(-2.5) = 0.8351.$
3. Since $P(X > x) = 1 - \Phi\left(\frac{x-5}{10}\right) = 0.05$, we must have $\Phi\left(\frac{x-5}{10}\right) = 0.95$. By using the normal distribution table, we can find $\frac{x-5}{10} \approx 1.64$. Thus, $x \approx 5 + (10)(1.64) = 21.4$.

Problem 4.

1. $P(|X-\mu| \leq 0.675\sigma) = P(-0.675\sigma \leq X-\mu \leq 0.675\sigma) =$ $\Phi(0.675) - \Phi(-0.675) \approx 0.5$
2. $P(\mu - c \leq X \leq \mu + c) = 0.95$ implies $P\left(\dfrac{X-\mu}{\sigma} \leq \dfrac{c}{\sigma}\right) = \Phi\left(\dfrac{c}{\sigma}\right) = 0.975$. Then we can find $\dfrac{c}{\sigma} \approx 1.96$. Thus, $c = 1.96\sigma$.

Problem 5.

By solving $P(X < 1) = 1 - e^{-\lambda} = 0.05$, we obtain $\lambda = -\ln 0.95 \approx 0.0513$.

Problem 6.

(a) Let p_1 be the probability that the component breaks down before 5 months, and let p_2 be the probability that the component breaks down before 50 months. Then,

$$p_1 = 1 - e^{-5\mu} \approx 1 - (1 - 0.0005) = 0.0005;$$
$$p_2 = 1 - e^{-50\mu} \approx 1 - (1 - 0.005) = 0.005.$$

(b) Let X_1 be the number of components breaking down before 5 months, and let X_2 be the number of components breaking down before 50 months. Then X_i has the binomial distribution with parameter $(p_i, 1000)$ for $i = 1, 2$. By using the Poisson approximation with $\lambda_1 = 1000p_1 = 0.5$ and $\lambda_2 = 1000p_2 = 5$, we can obtain

$$P(X_1 \leq 2) = e^{-\lambda_1}\left(1 + \lambda_1 + \frac{{\lambda_1}^2}{2}\right) \approx 0.986;$$

$$P(X_2 \leq 2) = e^{-\lambda_2}\left(1 + \lambda_2 + \frac{{\lambda_2}^2}{2}\right) \approx 0.125.$$

Problem 7.

Let X be the length of phone call in minutes.

(a) By solving $P(X \leq 20) = 1 - e^{-20\lambda} = 0.5$, we obtain $\lambda = \frac{\ln 0.5}{20} \approx 0.0347$.

(b) $P(10 \leq X \leq 20) = (1 - e^{-20\lambda}) - (1 - e^{-10\lambda}) \approx 0.207$.

Problem 8.

Since $Y = X_1^2 + X_2^2 + X_3^2 + X_4^2 + X_5^2 + X_6^2$ has a chi-square random variables with $n = 6$, we obtain $P(Y \leq 1.64) = 0.05$.

Problem 9.

(a) Recall that $a < 0$.

$$P(Y \leq t) = P(aX + b \leq t) = P(X \geq (t-b)/a) = \int_{(t-b)/a}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-}{2\sigma}\right.$$

$$= \int_{-\infty}^{t} \frac{1}{|a|\sigma\sqrt{2\pi}} \exp\left[-\frac{(y-a\mu-b)^2}{2(a\sigma)^2}\right] dy$$

by substituting $x = (y-b)/a$.

(b) The integrand $f(y)$ is a normal density with $(a\mu + b, (a\sigma)^2)$.

Problem 10.

For $t > 0$ we obtain

$$P(Y \leq t) = P(e^X \leq t) = P(X \leq \ln t) = \int_{-\infty}^{\ln t} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] dx$$
$$= \int_0^t \frac{1}{y\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln y-\mu)^2}{2\sigma^2}\right] dy$$

by substituting $x = \ln y$. Thus, the density is given by

$$f(y) = \frac{1}{y\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln y-\mu)^2}{2\sigma^2}\right], \quad y > 0.$$